DOCUMENT RESUME

ED 060 026 TM 001 109

AUTHOR	Shoemaker, David M.
TITLE	An Application of Item-Examinee Sampling to Scaling Attitudes.
INSTITUTION	Southwest Regional Educational Lab., Inglewood, Calif.
PUB DATE	Apr 72
NOTE	3p.; Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, Illinois, April 1972
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Attitudes; *Attitude Tests; *Item Sampling; Mathematical Applications; *Personality Assessment; *Personality Tests; Rating Scales; Response Style (Tests); Statistical Analysis; Stimuli; Test Interpretation; Transformations (Mathematics)
IDENTIFIERS	*Thurstons Model V Scaling Procedure

ABSTRACT

The post mortem item-examinee sampling investigation described herein explored the feasibility of using item-examinee sampling to estimate scale values denoting degree of affect toward stimuli when measured by the method of paired-comparisons. Results indicate clearly that such scale values can be approximated satisfactorily through item-examinee sampling. Defining one observation as the response made by one examinee to one item, the similarity between the estimated scale values and normative scale values increased generally with increases in the number of observations acquired by the sampling plan. (Author)

ED 060026

23.1

AN APPLICATION OF ITEM-EXAMINEE SAMPLING TO SCALING ATTITUDES

DAVID M. SHOEMAKER

Southwest Regional Laboratory for Educational Research and Development

The post mortem item-examinee sampling investigation described herein explored the feasibility of using item-examinee sampling to estimate scale values denoting degree of affect toward stimuli when measured by the method of paired-comparisons. Results indicate clearly that such scale values can be approximated satisfactorily through item-examinee sampling. Defining one observation as the response made by one examinee to one item, the similarity between the estimated scale values and normative scale values increased generally with increases in the number of observations acquired by the sampling plan.

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

TM 001 109

David M. Shoemaker
Southwest Regional Laboratory for Educational
Research and Development
11300 La Cienega Blvd.
Inglewood, California 90304

AN APPLICATION OF ITEM-EXAMINEE SAMPLING TO SCALING ATTITUDES

DAVID M. SHOEMAKER

Southwest Regional Laboratory for Educational
Research and Development

Suppose that you have a number of stimuli (e.g., physical objects, statements about some issue, job applicants, employees) and you want not only to rank them according to some underlying attribute but to assign them scale values. One procedure is that of asking a number of judges (i.e., examinees, subjects) to consider all possible pairs of stimuli and decide which one of each pair has more of the underlying attribute. The results of these paired-comparisons are then analyzed further and the desired scale values determined. One popular method (and the one used in this investigation) is Thurstone's Model V scaling procedure.

When several stimuli are being scaled a problem arises in that the number of possible paired-comparisons becomes large. Specifically, the number of paired-comparisons is equal to $s(s - 1)/2$ where s is the number of stimuli being compared. The investigation described herein was concerned with determining (a) the extent to which scale values derived from giving all pairings to all judges could be estimated through giving subsets of pairings selected randomly to random samples of judges (a procedure referred to frequently in the literature as item-examinee sampling), and (b) the standard errors of estimate associated with each sampling plan.

METHOD

The research design was one of post mortem item-examinee sampling: given a data base, various item-examinee samples (samples of paired-comparisons and judges) are selected randomly from this base and used to estimate parameters of the data base from which they have been sampled. The data base used in this investigation consisted of responses made by 407 primary grade students to a 15-item test designed to scale degree of affect to 6 stimuli by the method of paired-comaprisons.

In the method of paired-comparisons, the preliminary analysis of the data involves the construction of an F-matrix and subsequent P-matrix. If the s stimuli are numbered consecutively from 1 to s, the F-matrix is an s by s matrix with cell entries denoting the frequency with which the column stimulus was judged more favorable than the row stimulus. Dividing each entry in the F-matrix by the total number of examinees produces the corresponding entry in the P-matrix labeled appropriately as the proportion

of examinees selecting the column stimulus over the row stimulus. In this investigation, the proportion of times stimulus i was preferred over stimulus j among those examinees comparing i and j was used to estimate the ij cell in the P-matrix.

In an attempt to determine both the relative merits of individual sampling plans and general procedures for estimating scale values through item-examinee sampling, 10 individual sampling plans were examined. Incorporated within the sampling plans were selected variations in the number of subtests, the number of items per subtest, and the number of examinees responding to each subtest. The results of each sampling plan were replicated 25 times to approximate the sampling errors associated with each sampling plan.

Each replication of each sampling plan produced one set of 6 scale values. These scale values are estimates of the 6 scale values obtained by testing 407 examinees over 15 items. The statistic judged most appropriate for quantifying the degree of similarity between the sample scale values and the normative scale values was the generalized distance function. Because the range for each set of 6 scale values changed from replication to replication, all scale values within each set were multiplied by the inverse of the largest scale value before computing the generalized distance. Without this transformation, the pooling of the generalized distances among replications would have been inappropriate.

RESULTS AND DISCUSSION

Defining one observation as the response made by one examinee to one item, the results support the conclusion that, as the number of observations acquired by the sampling plan increases, the estimated scale values approximate more closely the scale values obtained by testing all examinees over all items. In addition, the standard error of estimate decreases generally with increases in the number of observations. The results of this investigation suggest strongly that scale values determined by the method of paired-comparisons can be approximated satisfactorily by item-examinee sampling procedures.

Although item-examinee sampling procedures were developed originally to estimate parameters for achievement tests, the extension of such procedures to other psychometric instruments appears to be a fruitful line of research. One obvious extension is to questionnaire and survey results.